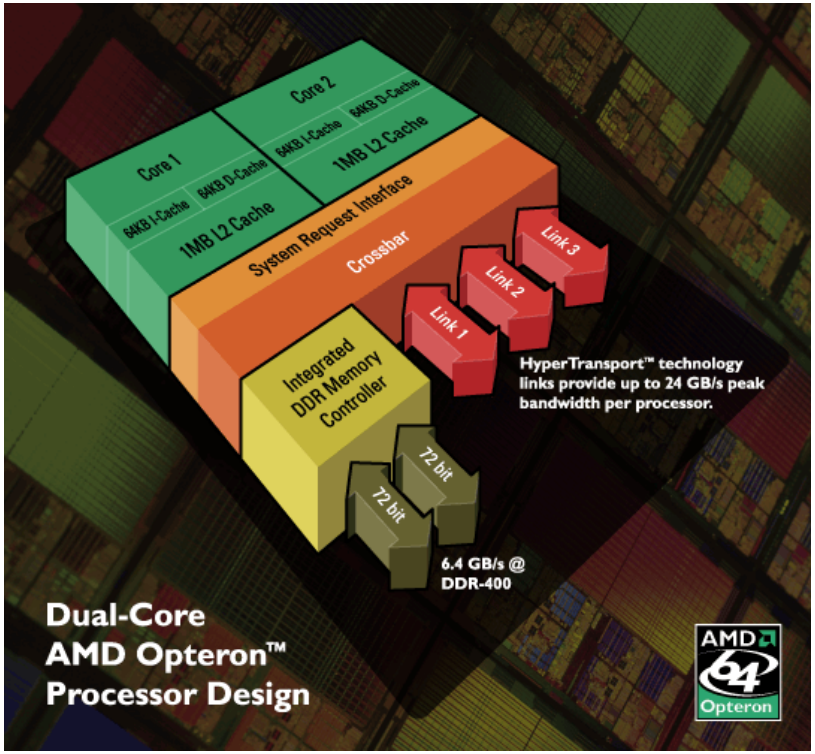
The AMD Opteron™ CMP NorthBridge Architecture: Now and in the Future

Pat Conway & Bill Hughes

August, 2006

AMD Opteron™ – The Industry's First Native Dual-Core 64-bit x86 Processor



Integration:

- Two 64-bit CPU cores
- 2MB L2 cache
- On-chip Router & Memory Controller

Bandwidth:

- Dual channel DDR (128-bit) memory bus
- 3 HyperTransport™ (HT) links (16-bit each x 2 GT/sec x 2)

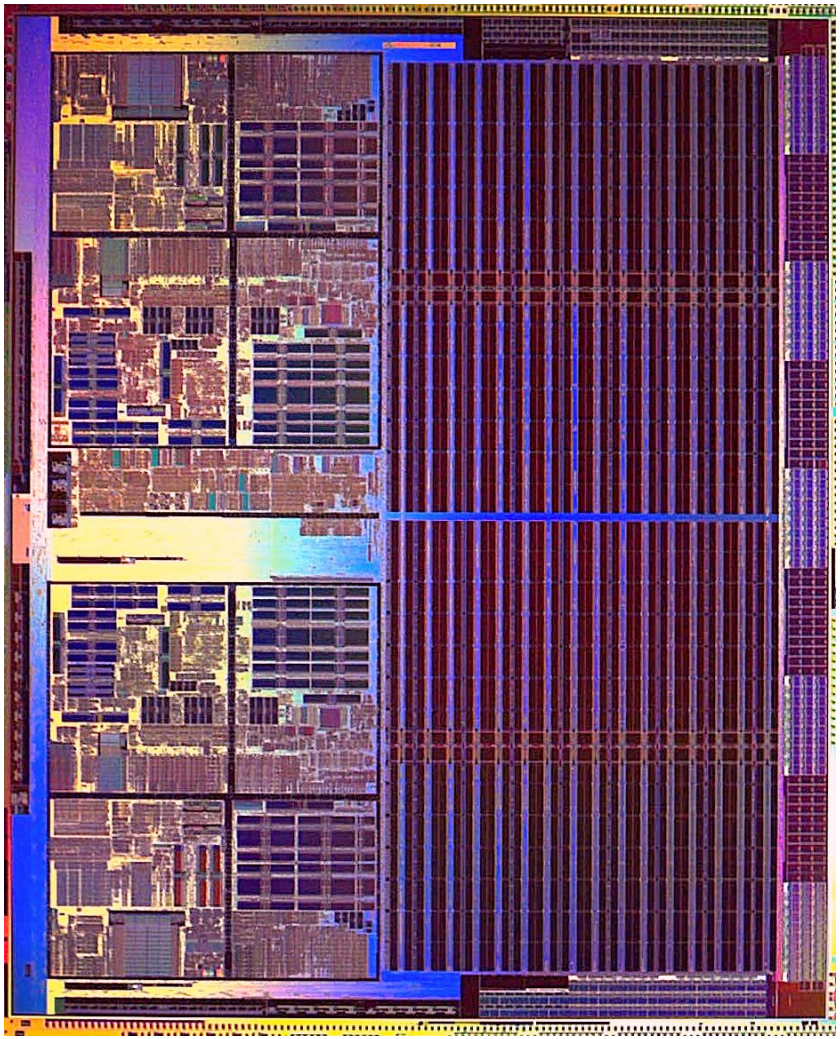
Usability and Scalability:

- Socket compatible: *Platform and TDP!*
- Glueless SMP up to 8 sockets
- Memory capacity & BW scale w/ CPUs

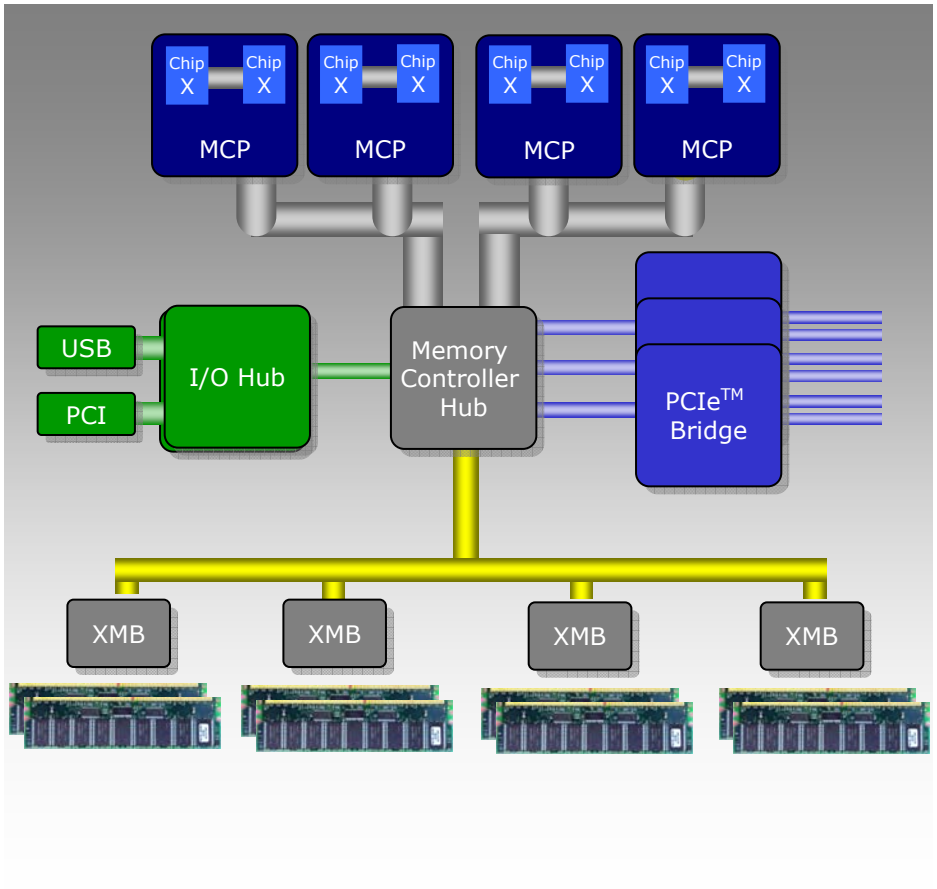
Power Efficiency:

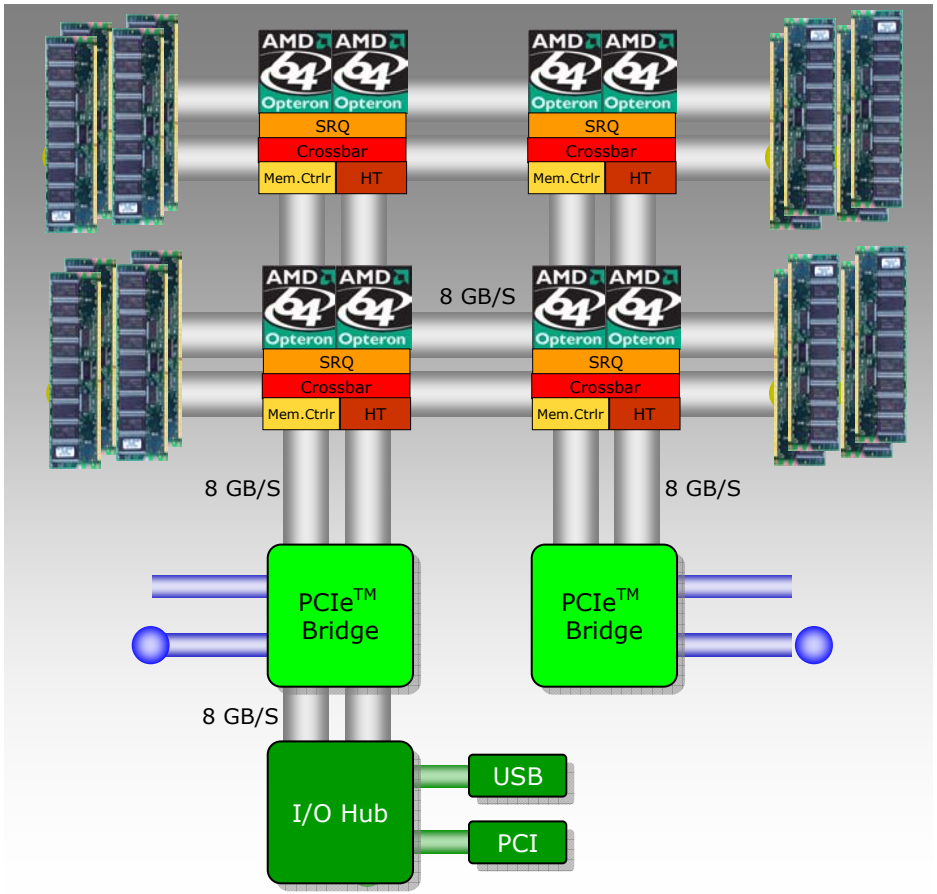
- AMD PowerNow!™ Technology with optimized power management
- Industry-leading system level power efficiency

AMD Opteron™ – The Industry's First Native Dual-Core 64-bit x86 Processor



A Clean Break with the Past





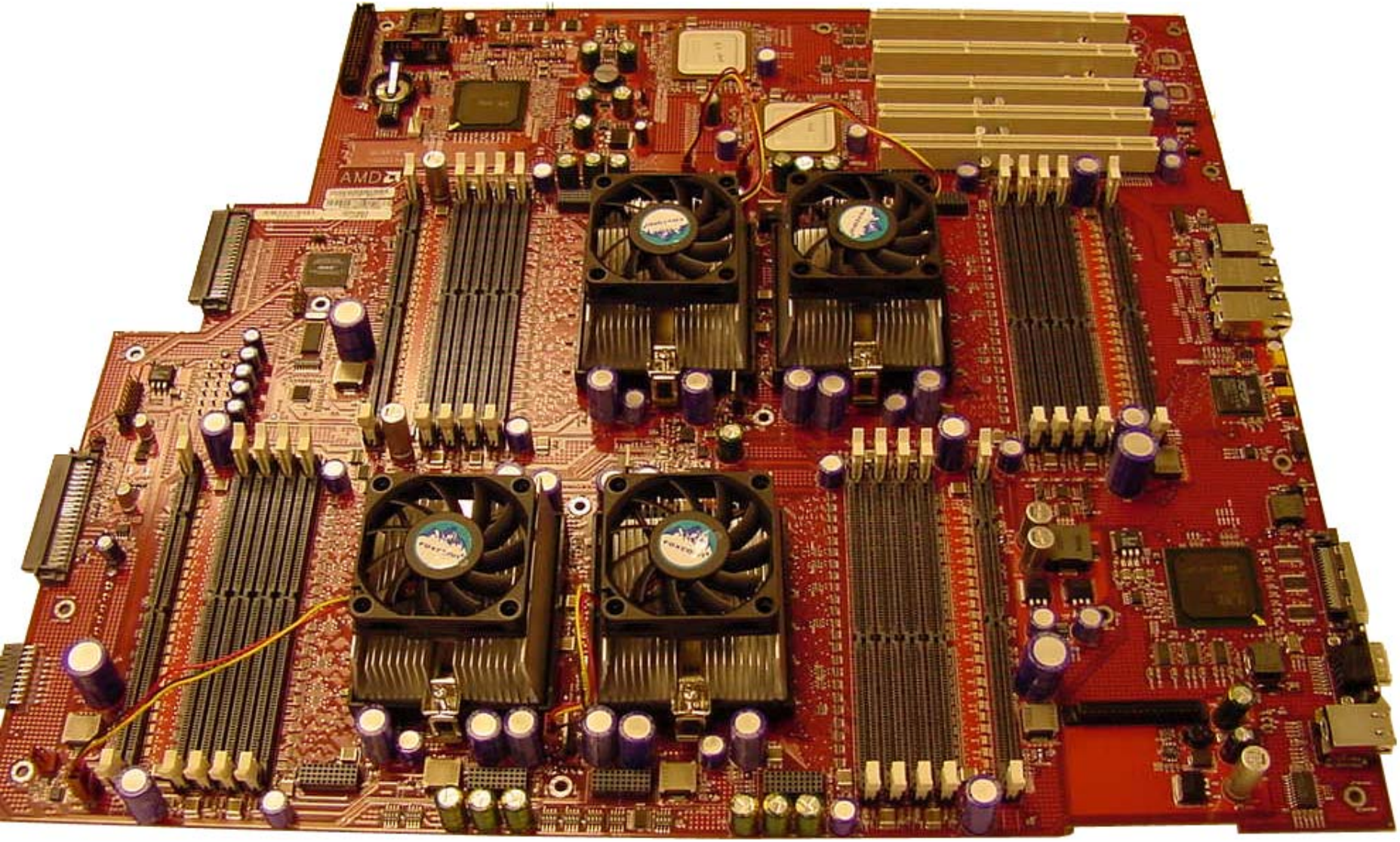
Legacy x86 Architecture

- 20-year old traditional front-side bus (FSB) architecture
- CPUs, Memory, I/O all share a bus
- Major bottleneck to performance
- Faster CPUs or more cores ≠ performance

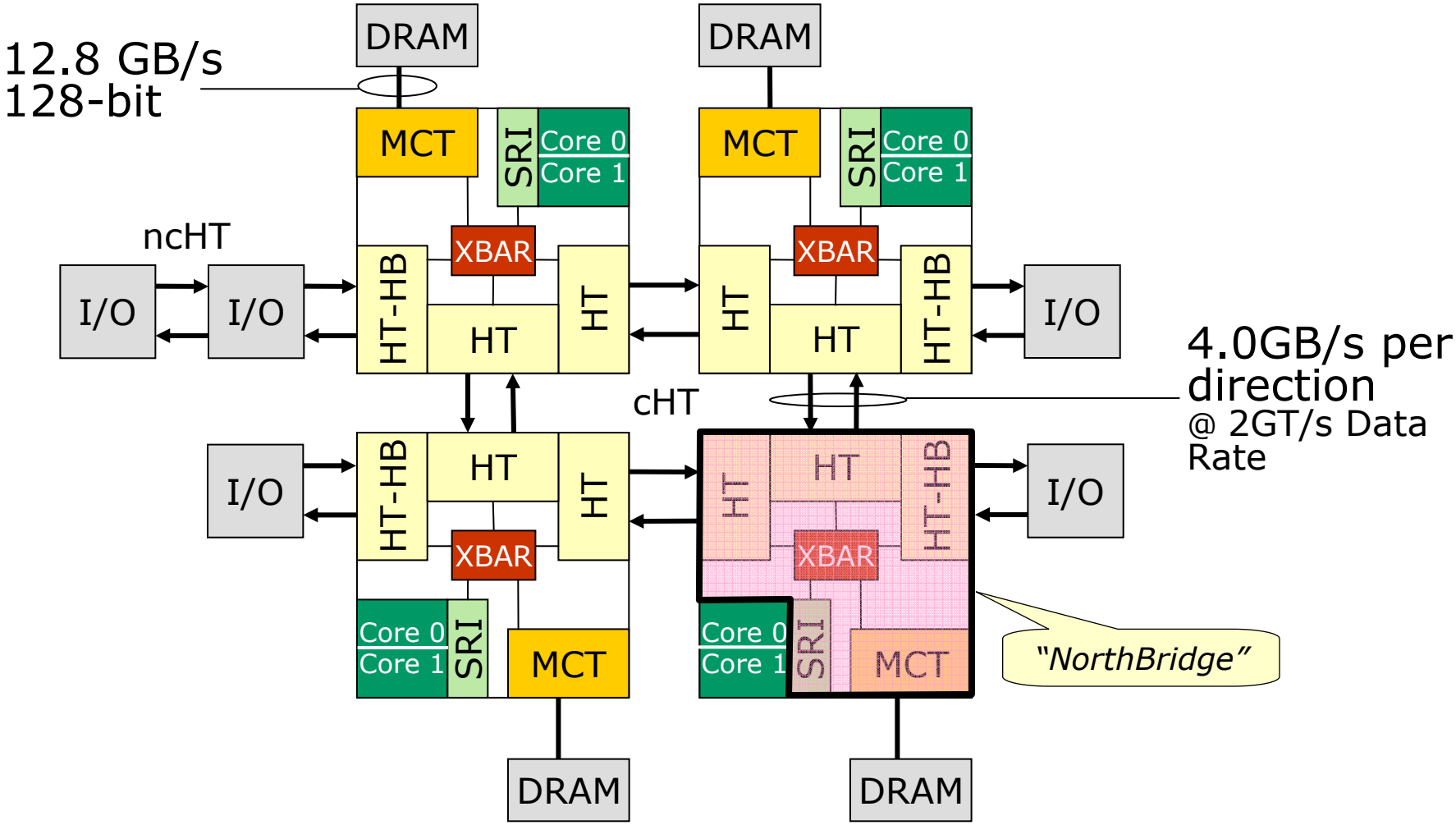
AMD64's Direct Connect Architecture

- Industry-standard technology
- Direct Connect Architecture reduces FSB bottlenecks
- HyperTransport™ interconnect offers scalable high bandwidth and low latency
- 4 memory controllers – increases memory capacity and bandwidth

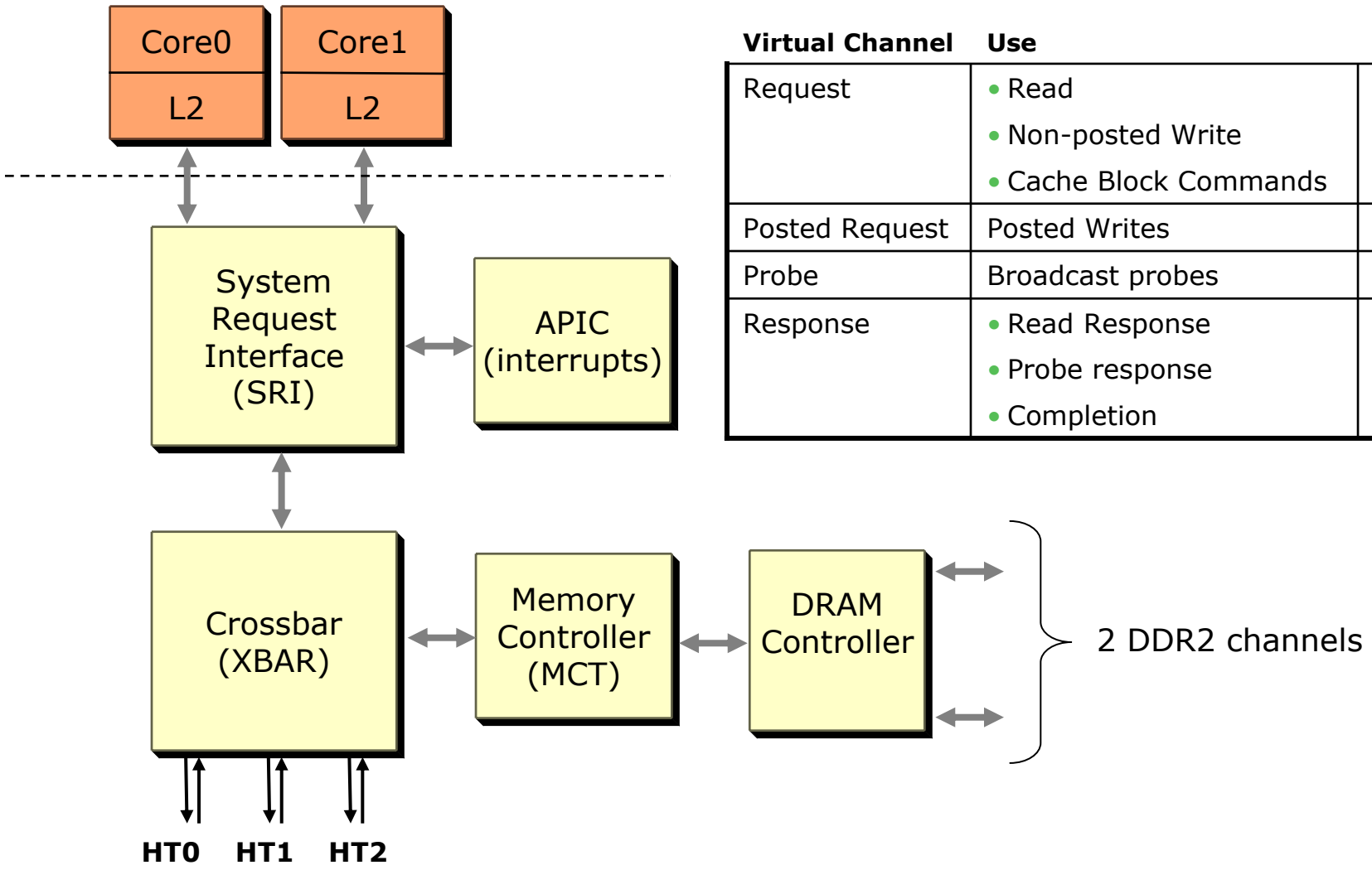
4P System — Board Layout



System Overview

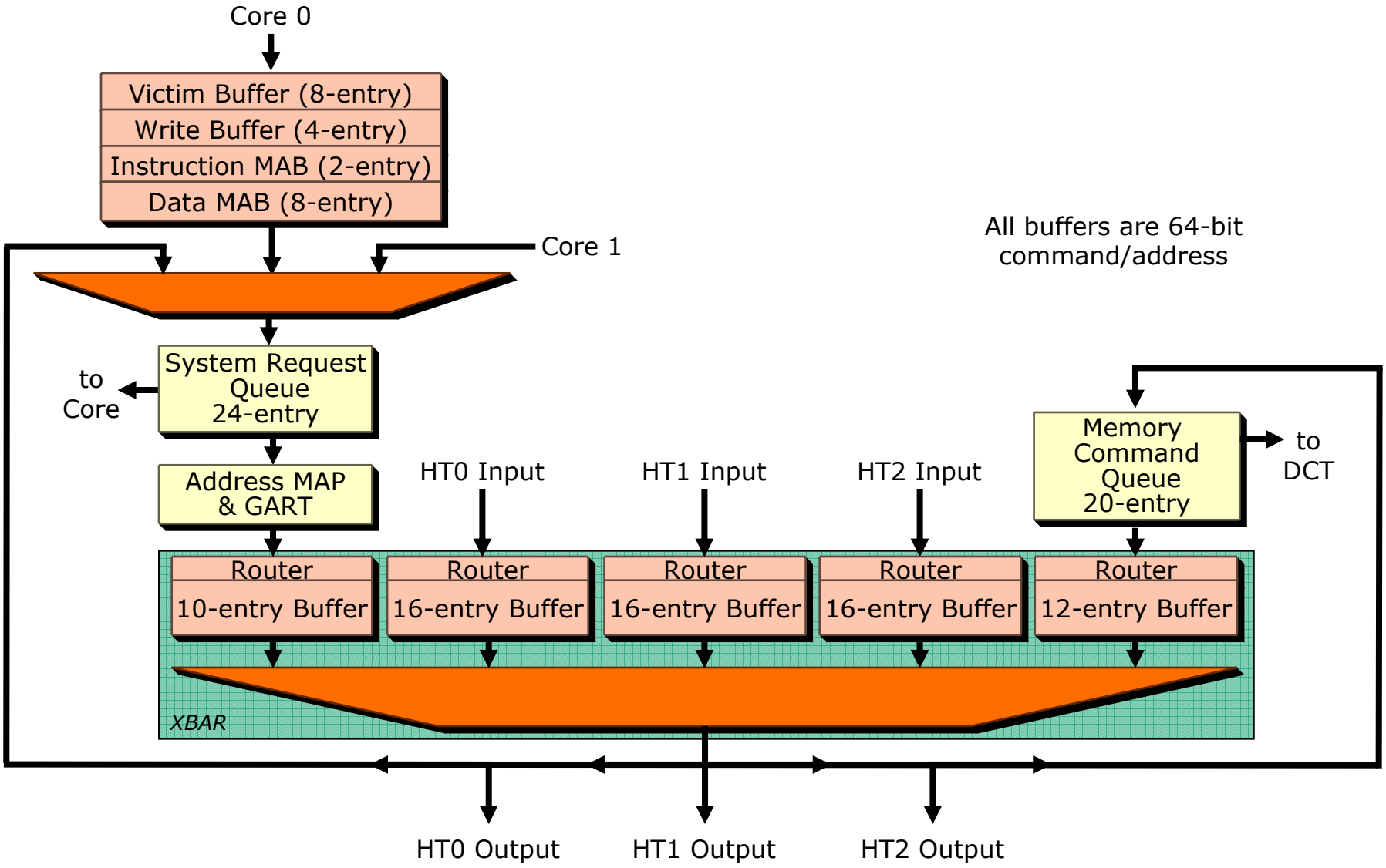


Northbridge Microarchitecture Overview

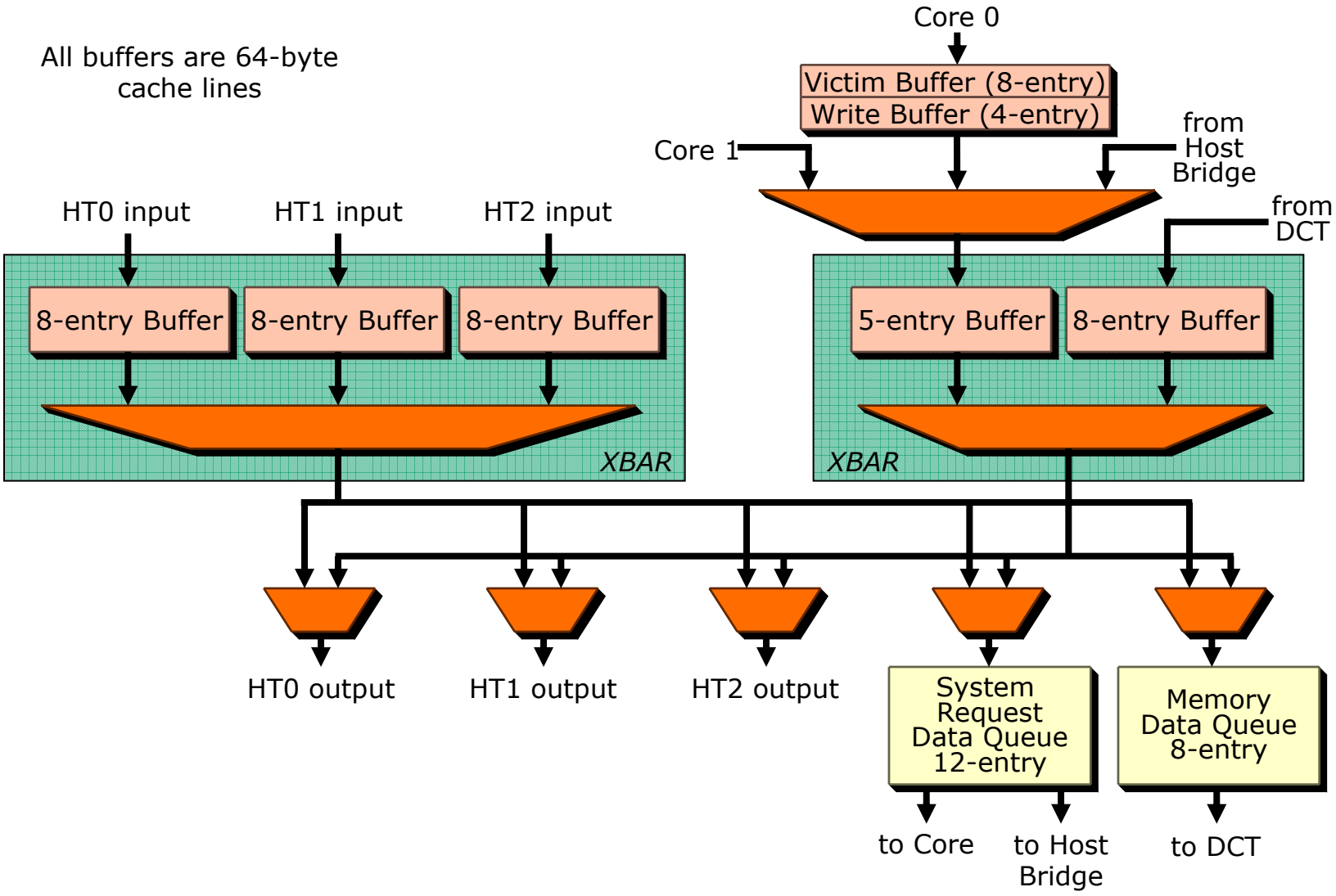


Virtual Channel	Use	Data
Request	• Read • Non-posted Write • Cache Block Commands	Y
Posted Request	Posted Writes	Y
Probe	Broadcast probes	N
Response	• Read Response • Probe response • Completion	Y

Northbridge Command Flow

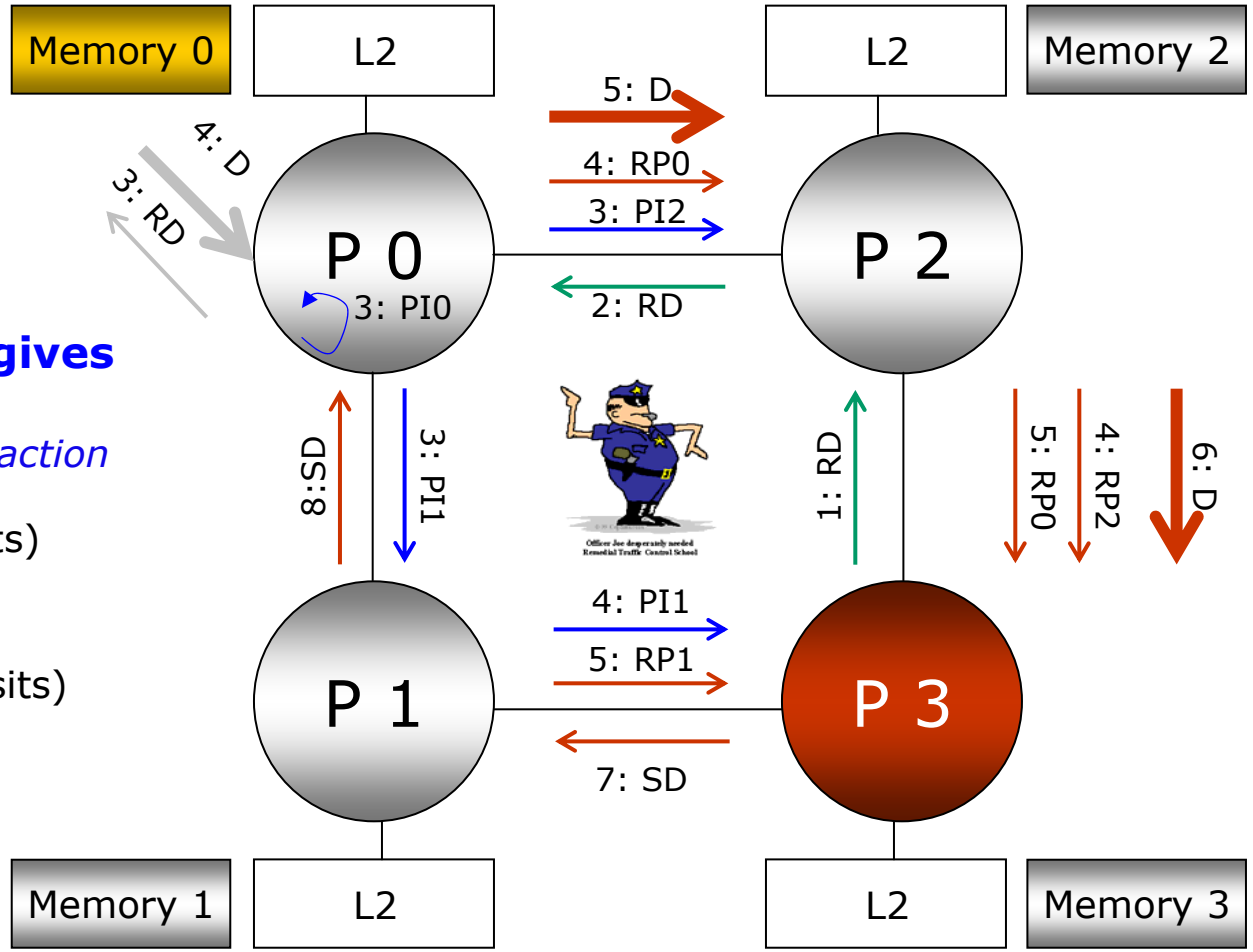


Northbridge Data Flow



Lessons Learned #1

Allocation of XBAR Command buffer across Virtual Channels can have big impact on performance



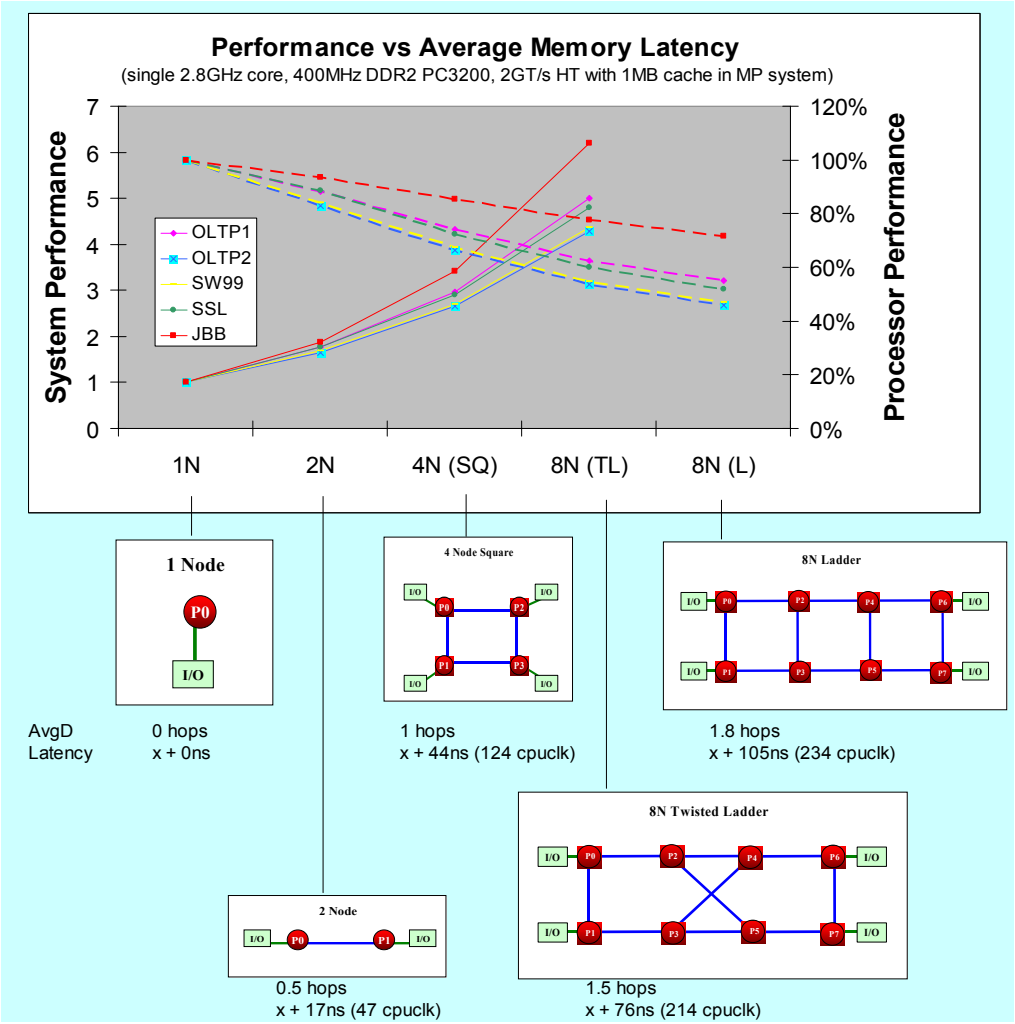
MP traffic analysis gives the best allocation

e.g. Opteron Read Transaction

- Request (2 visits)
- Probe (3 visits)
- Response (8 visits)

Lessons Learned #2

Memory Latency is the Key to Application Performance!

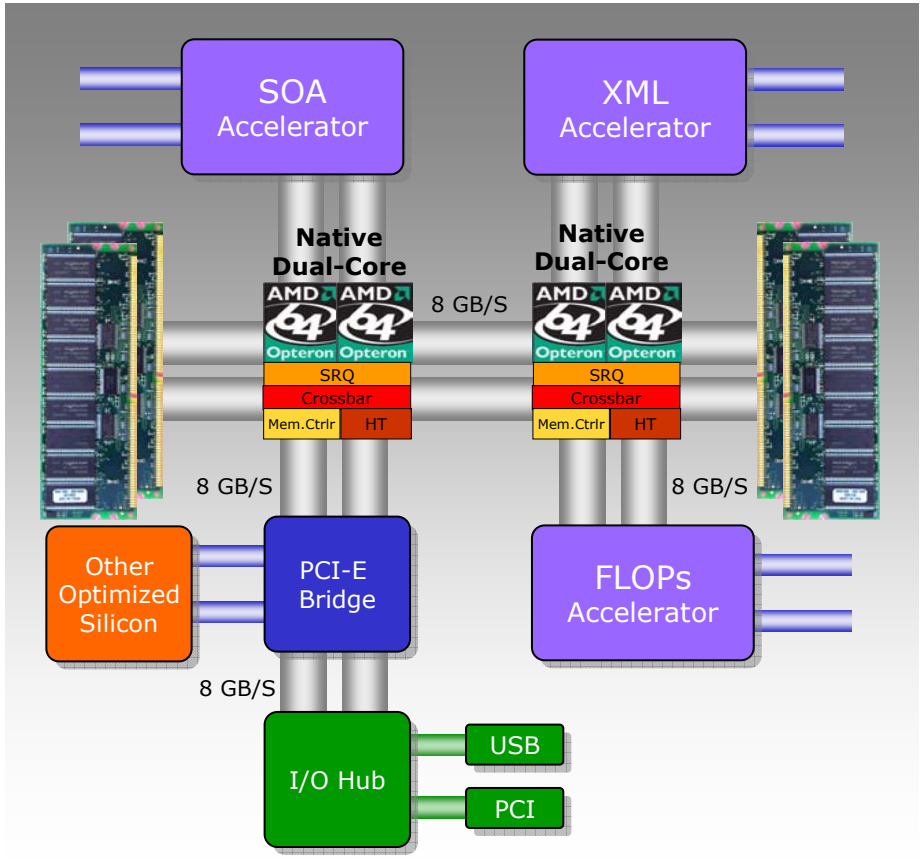


Looking Forward

HyperTransport™-based Accelerators

Imagine it, Build it

- Open platform for system builders ("Torrenza")
  - 3rd Party Accelerators
  - Media
  - FLOPs
  - XML
  - SOA
- AMD Opteron™ Socket or HTX slot
- HyperTransport interface is an open standard see hypertransport.org
- Coherent HyperTransport interface available if the accelerator caches system memory (under license)

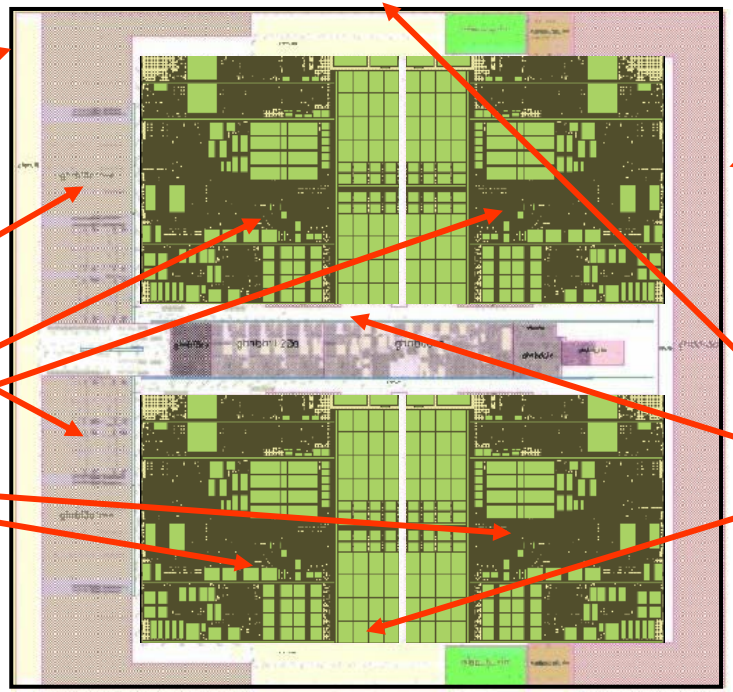


AMD's Next Generation Processor Technology

Native quad core die

Expandable shared L3 cache

IPC enhanced CPU cores



Ideal for 65nm SOI and beyond

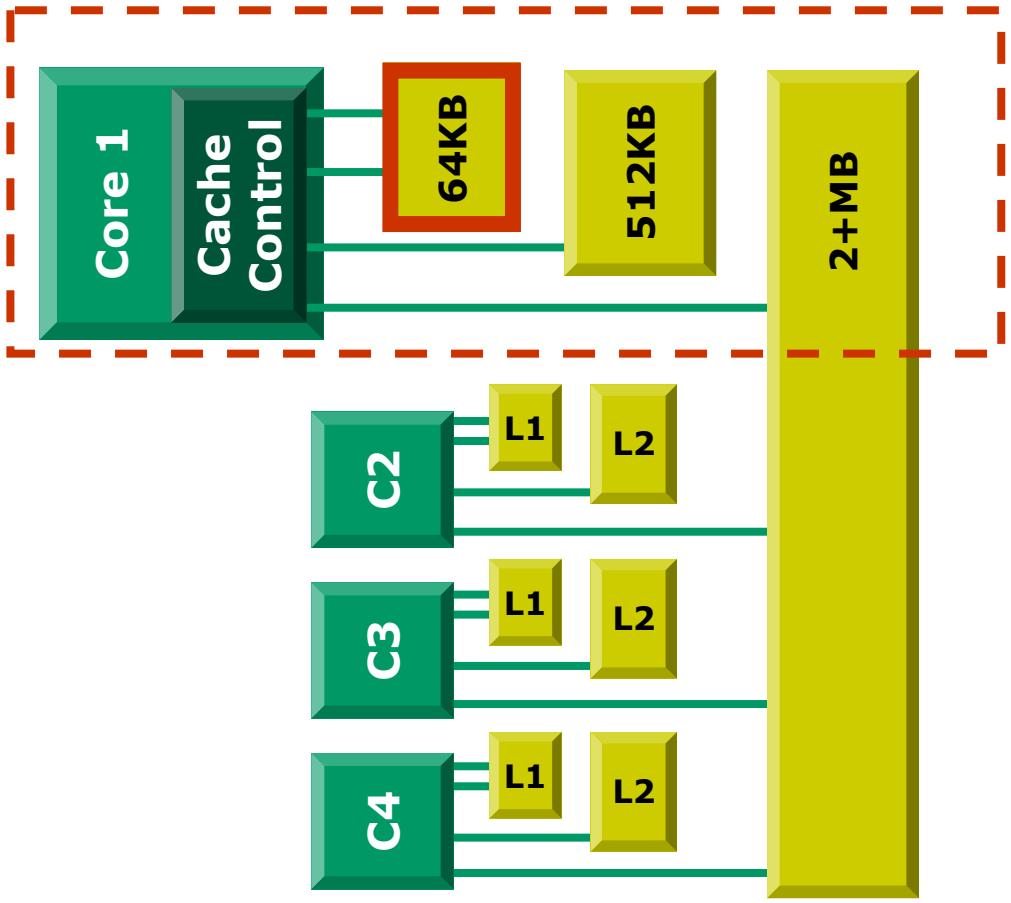
Enhanced Direct Connect Architecture and Northbridge

- 32B instruction fetch
- Improved branch prediction
- Out-of-order load execution
- Up to 4 DP FLOPS/cycle
- Dual 128-bit SSE dataflow
- Dual 128-bit loads per cycle
- Bit Manipulation extensions (LZCNT/POPCNT)
- SSE extensions (EXTRQ/INSERTQ, MOVNTSD/MOVNTSS)

- Four ungangable x16 HyperTransport™ links (up to 5.2GT/sec)
- Enhanced crossbar
- Next-generation memory support
- FBDIMM *when appropriate*
- Enhanced power management and RAS

Balanced, Highly Efficient Cache Structure

Efficient memory handling reduces the need for "brute force" cache sizes

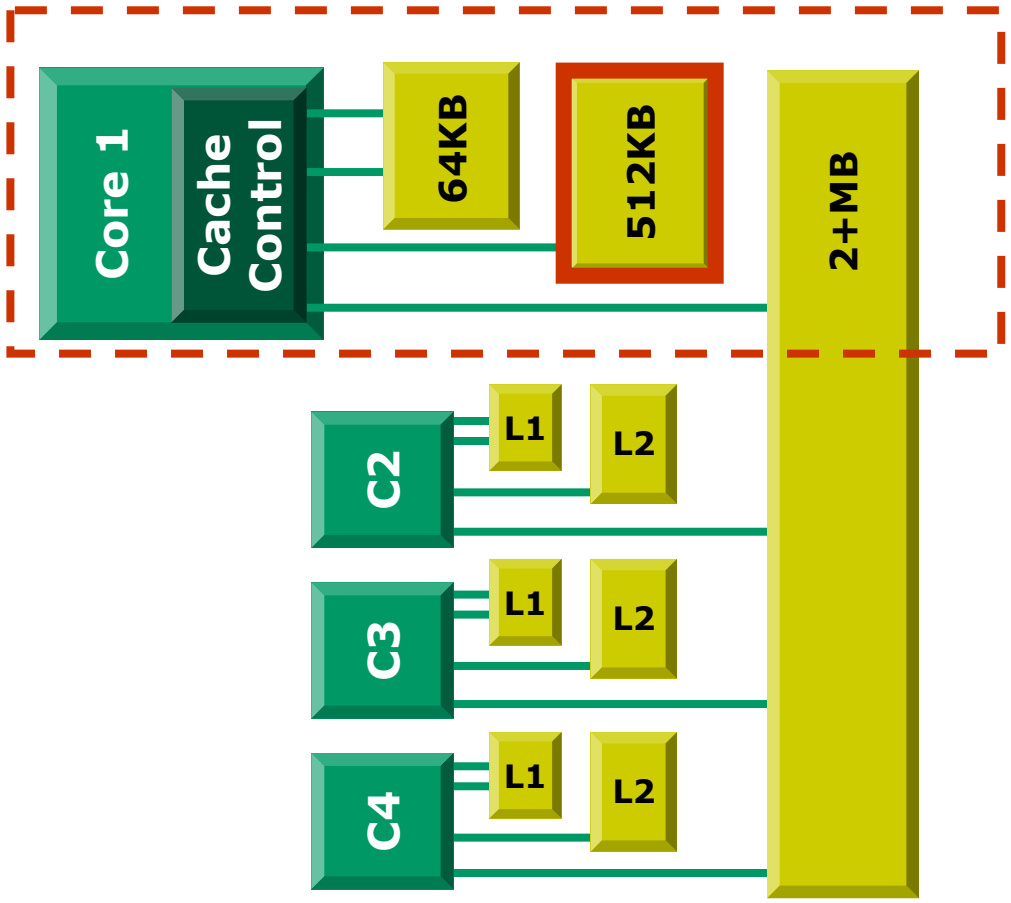


Dedicated L1

- Locality keeps most critical data in the L1 cache
- Low latency
- 2 128 bit data paths
- 2 loads per cycle

Balanced, Highly Efficient Cache Structure

Efficient memory handling reduces the need for "brute force" cache sizes

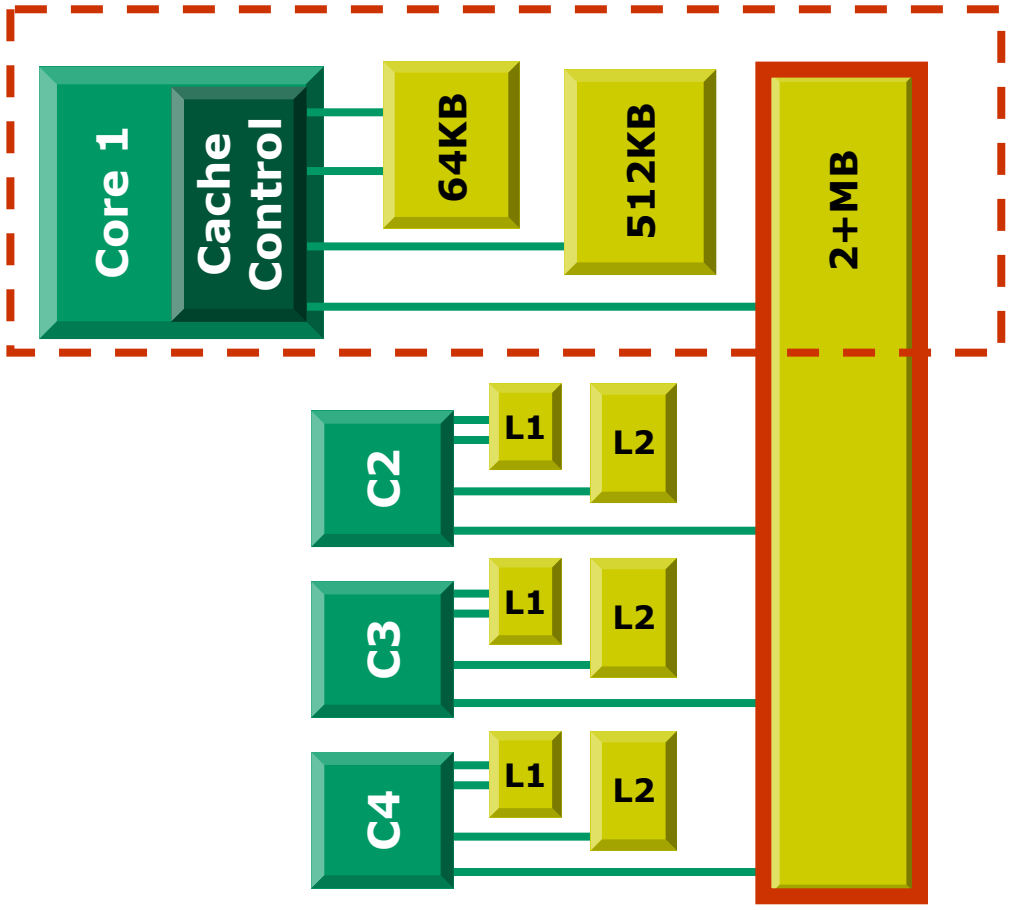


Dedicated L2

- Sized to accommodate the majority of working sets today
- Dedicated to help eliminate conflicts common in shared caches

Balanced, Highly Efficient Cache Structure

Efficient memory handling reduces the need for "brute force" cache sizes

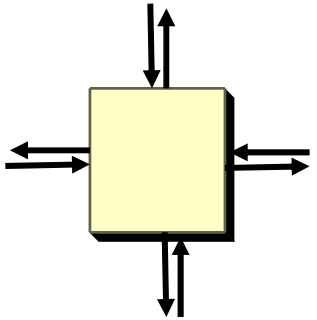


Shared L3 – *Coming Soon*

- Allocation policy which optimizes movement, placement and replication of data for multi-core
- Ready for expansion

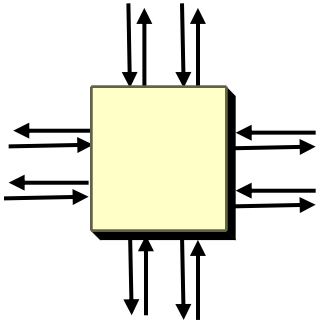
Additional HyperTransport™ Ports

- Enable Fully Connected 4 Node (four x16 HT) and 8 Node (eight x8 HT)
- Reduced network diameter
  - Fewer hops to memory
- Increased Coherent Bandwidth
  - more links
  - cHT packets visit fewer links
  - HyperTransport3
- Benefits
  - Low latency because of lower diameter
  - Evenly balanced utilization of HyperTransport links
  - Low queuing delays



Four x16 HT

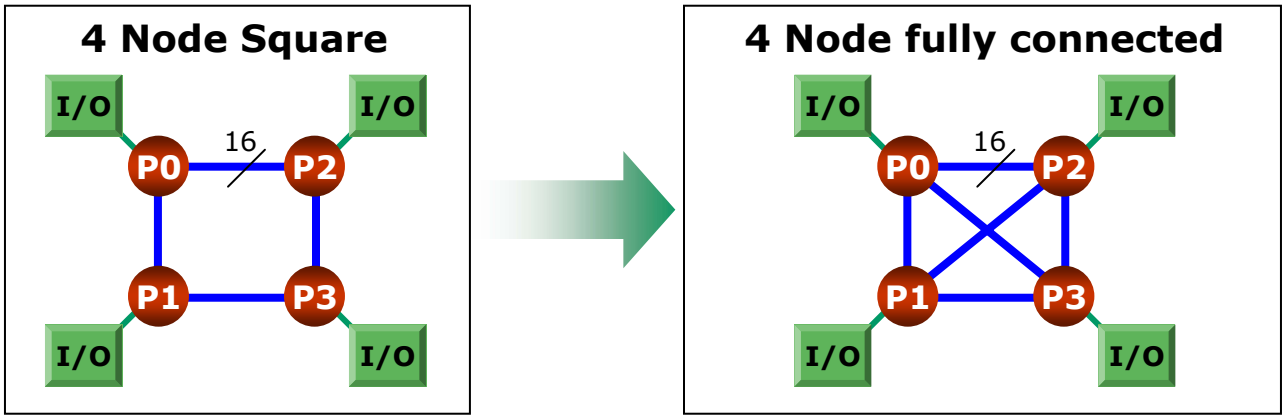
OR



Eight x8 HT

Low latency under load

4 Node Performance



4N SQ (2GT/s HyperTransport)

Diam 2 Avg Diam 1.00

XFIRE BW 14.9GB/s

+ 2 EXTRA LINKS

4N FC (2GT/s HyperTransport)

Diam 1 Avg Diam 0.75

XFIRE BW 29.9GB/s

(2X)

W/ HYPERTRANSPORT3

4N FC (4.4GT/s HyperTransport3)

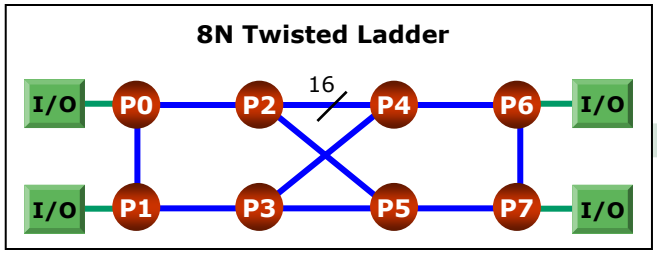
Diam 1 Avg Diam 0.75

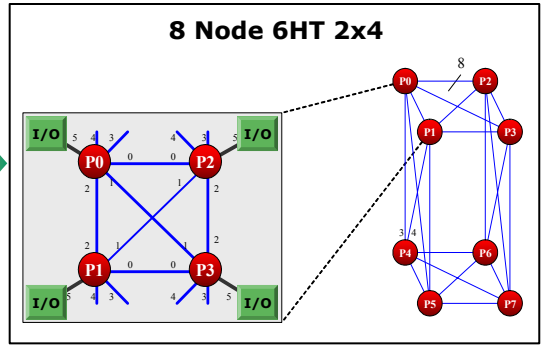
XFIRE BW 65.8GB/s

(4X)

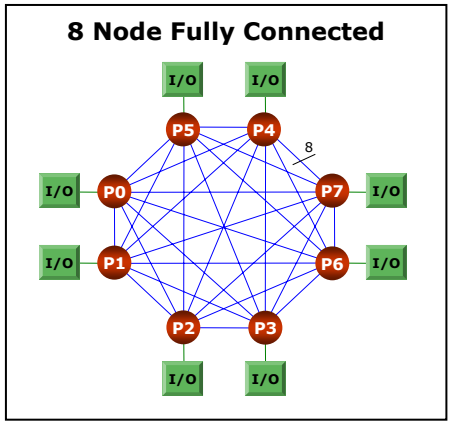
XFIRE ("crossfire") BW is the ***link-limited*** all-to-all communication bandwidth (data only)

8 Node Performance





OR



8N TL (2GT/s HyperTransport)

Diam 3 Avg Diam 1.62

XFIRE BW 15.2GB/s

8N 2x4 (4.4GT/s HyperTransport3)

Diam 2 Avg Diam 1.12

XFIRE BW 72.2GB/s

(5X)

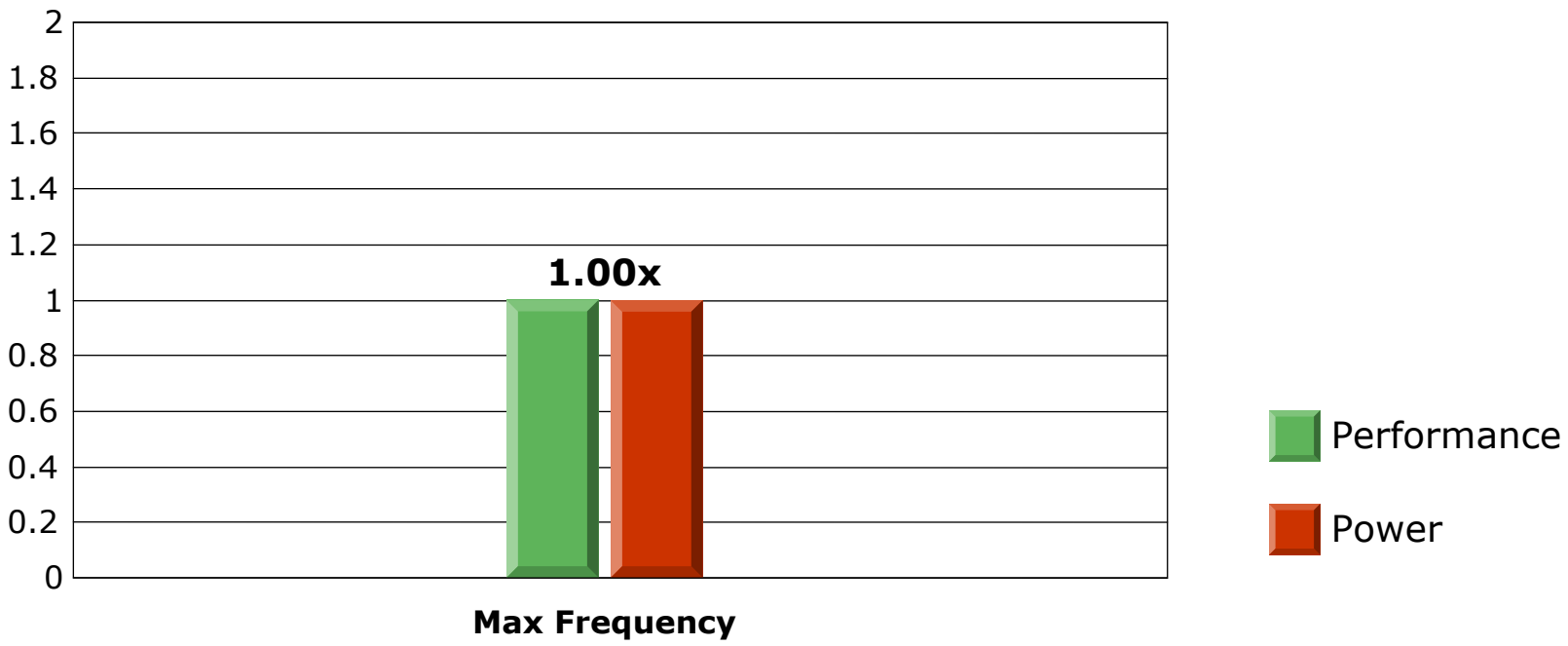
8N FC (4.4GT/s HyperTransport3)

Diam 1 Avg Diam 0.88

XFIRE BW 94.4GB/s

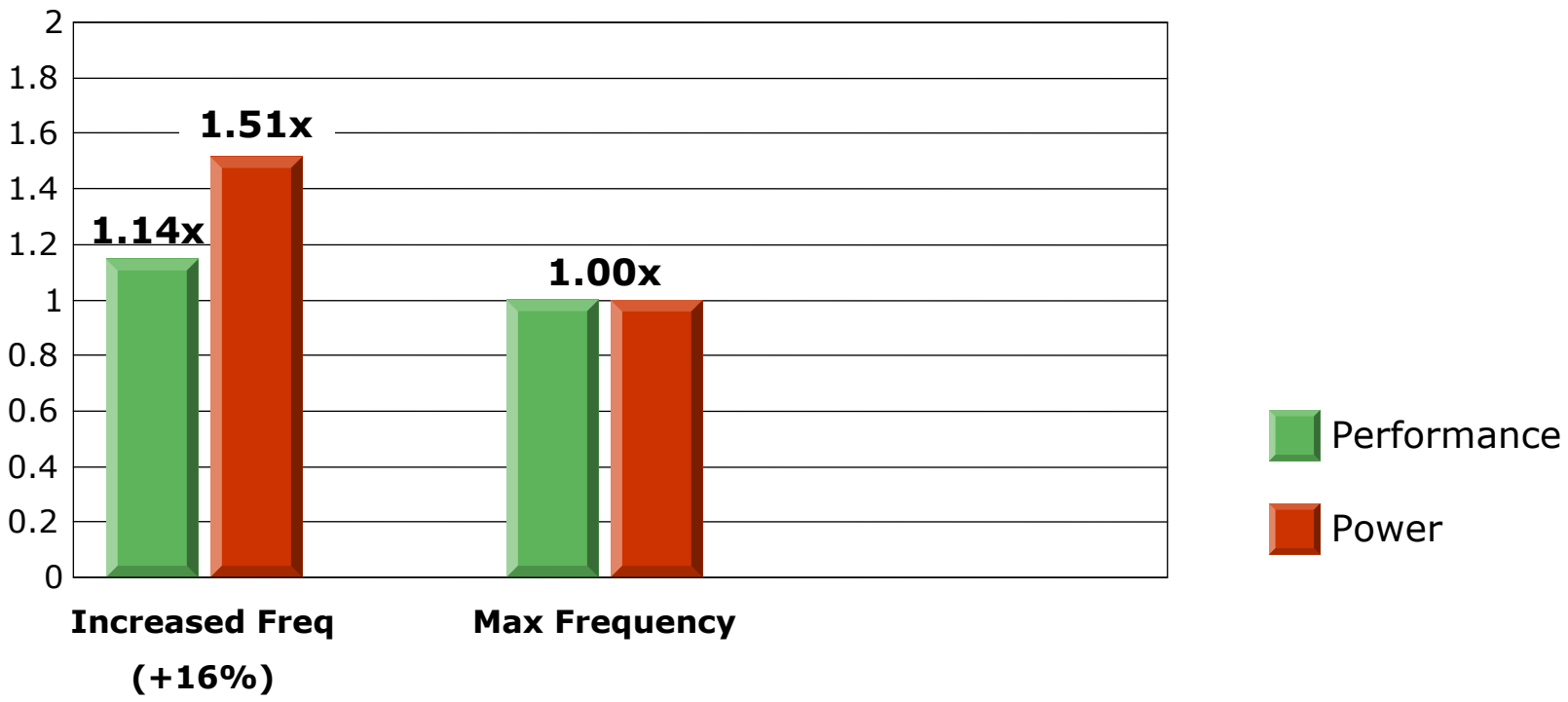
(6X)

Why Quad-Core?



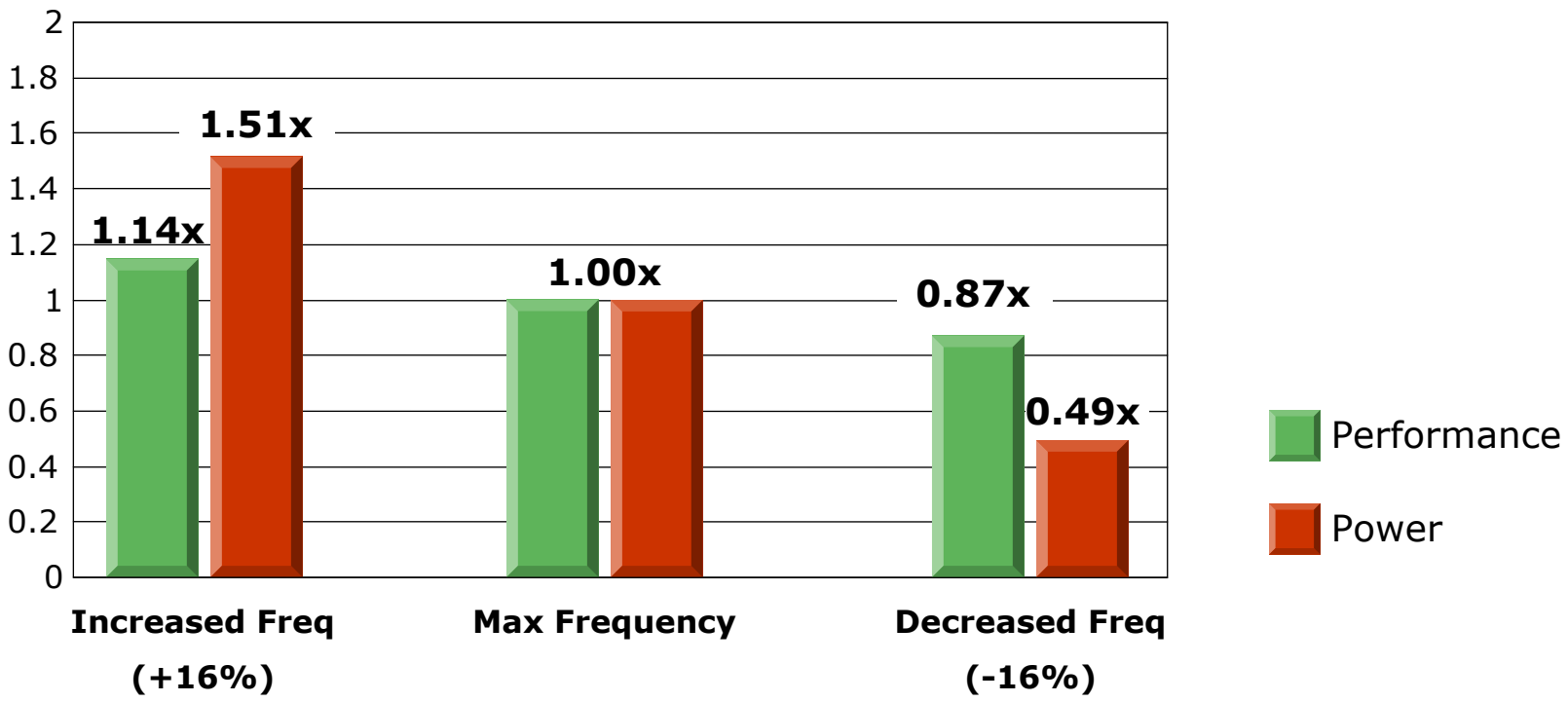
Baseline is 2 Node x 2 Core blade running OLTP

Increasing Frequency



Baseline is 2 Node x 2 Core blade running OLTP

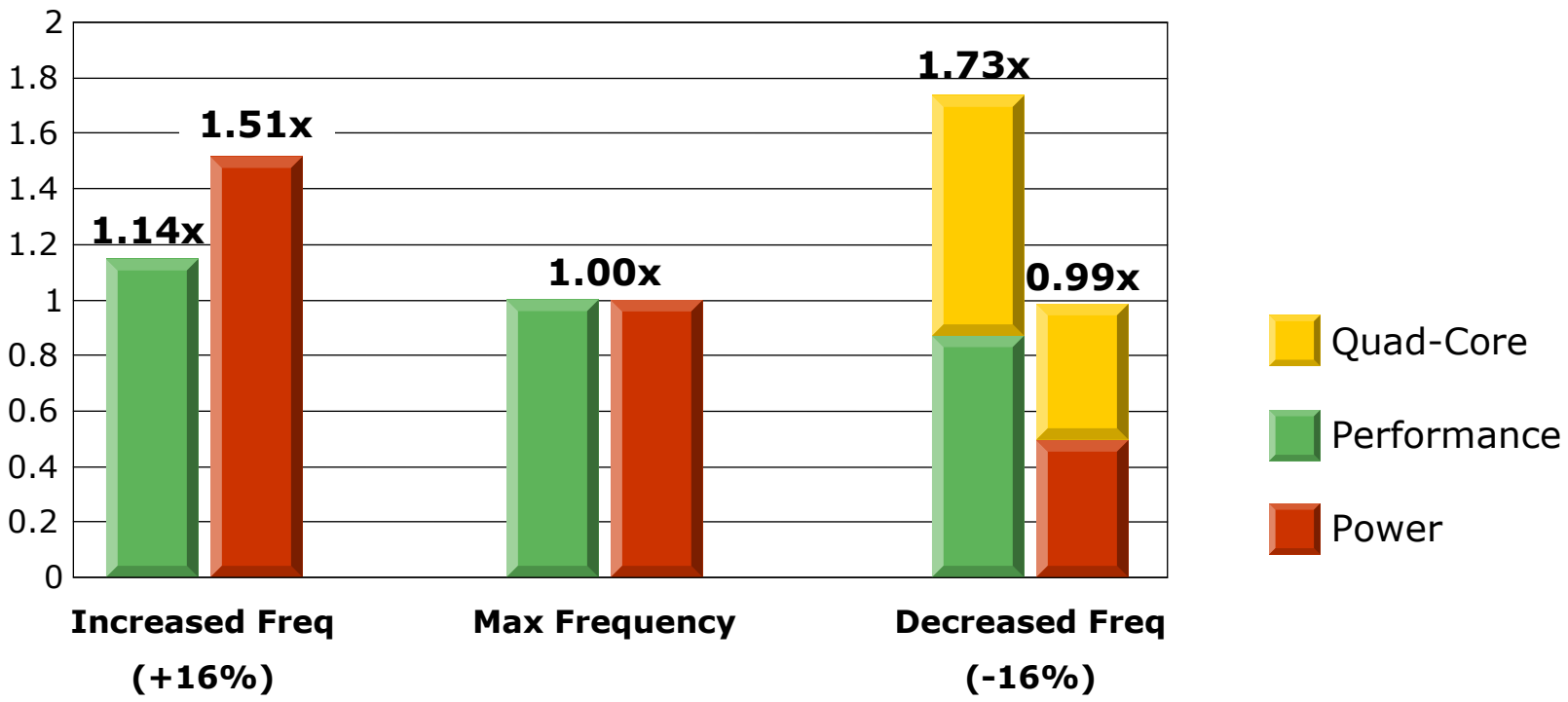
Decreasing Frequency



Baseline is 2 Node x 2 Core blade running OLTP

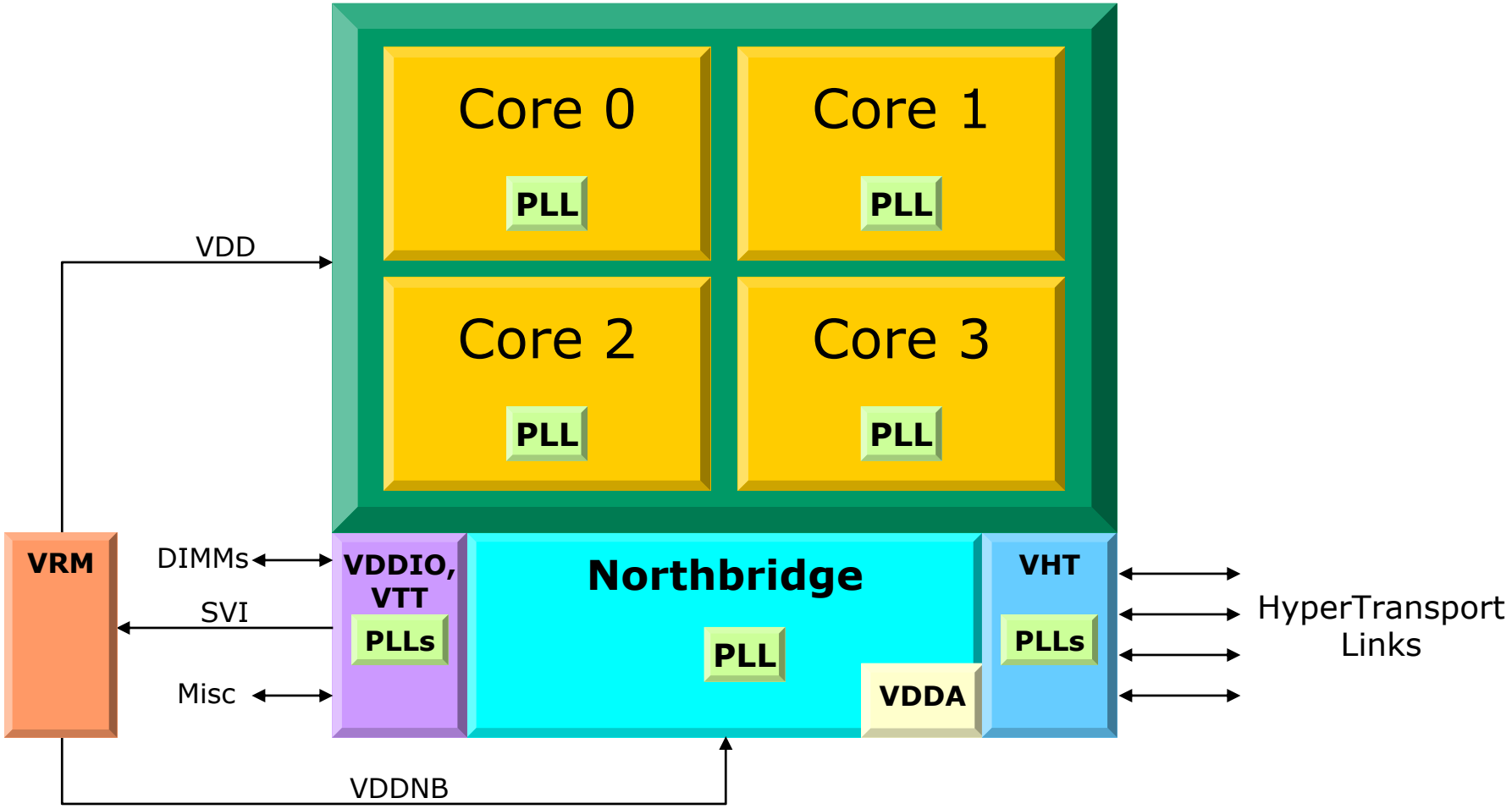
Quad-Core

Higher Performance within a Fixed Power Budget



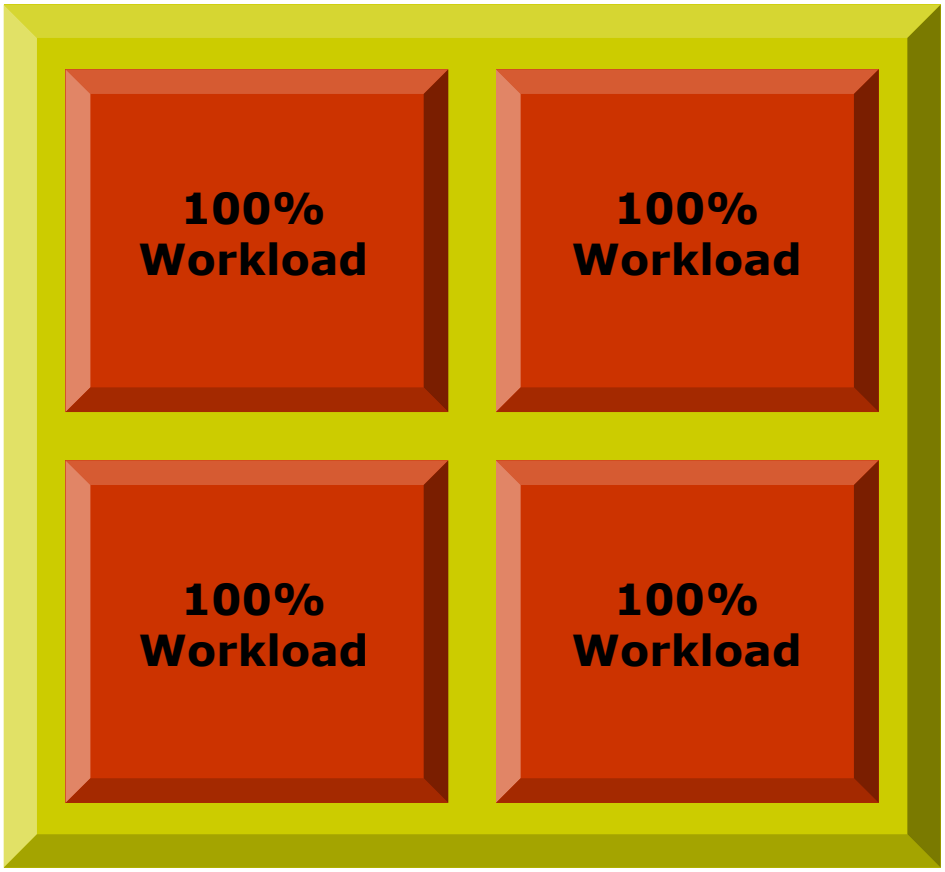
Baseline is 2 Node x 2 Core blade running OLTP

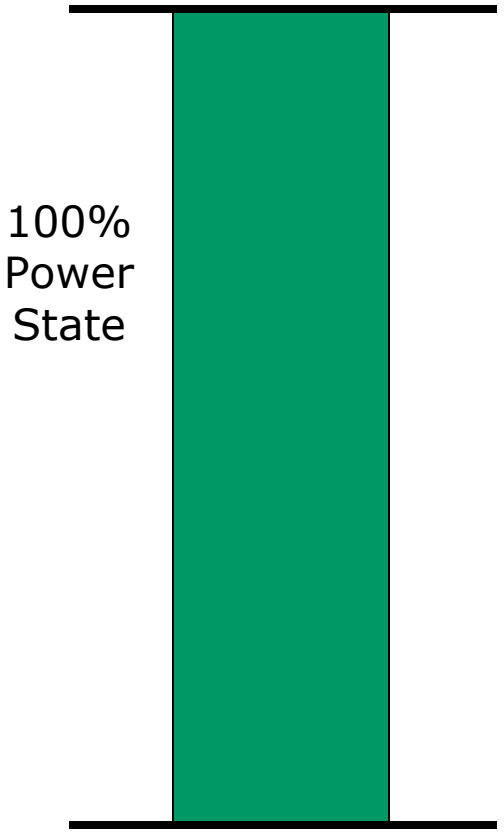
Clock and Power Planes



DICE: Dynamic Independent Core Engagement

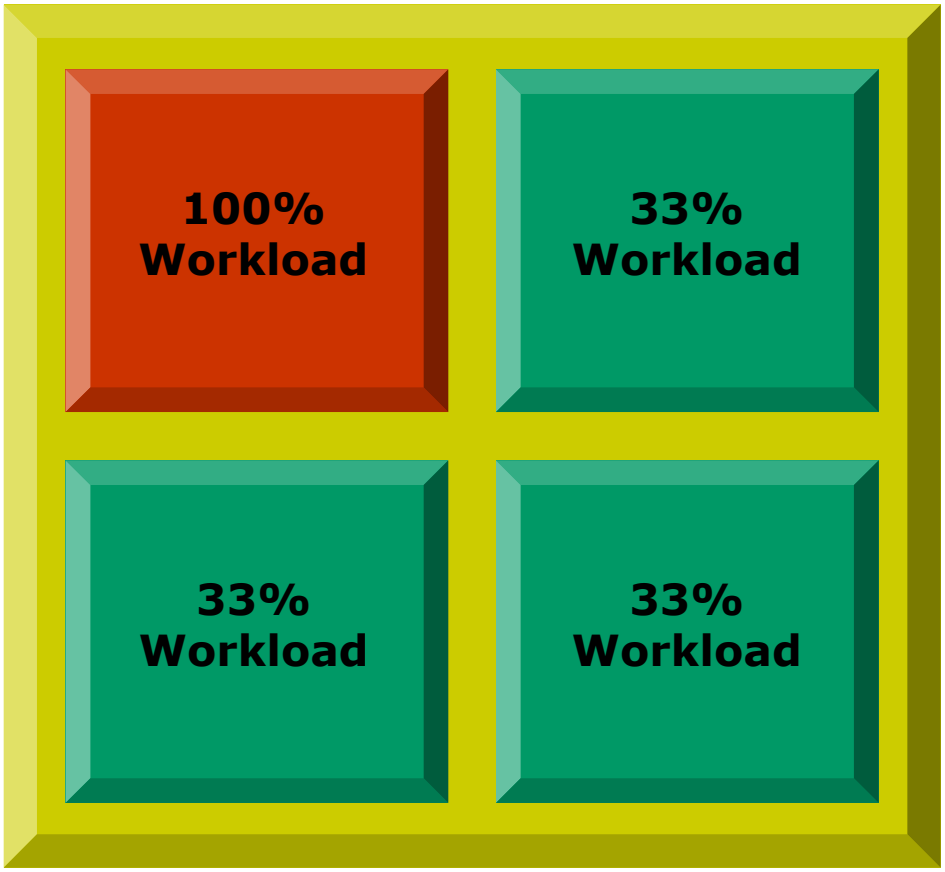
Ability to dynamically and individually adjust core frequencies to improve power efficiency

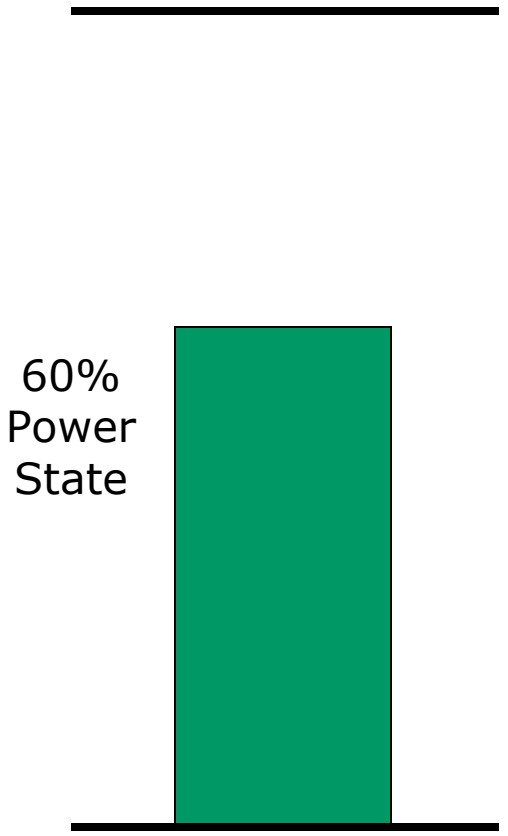




DICE: Dynamic Independent Core Engagement

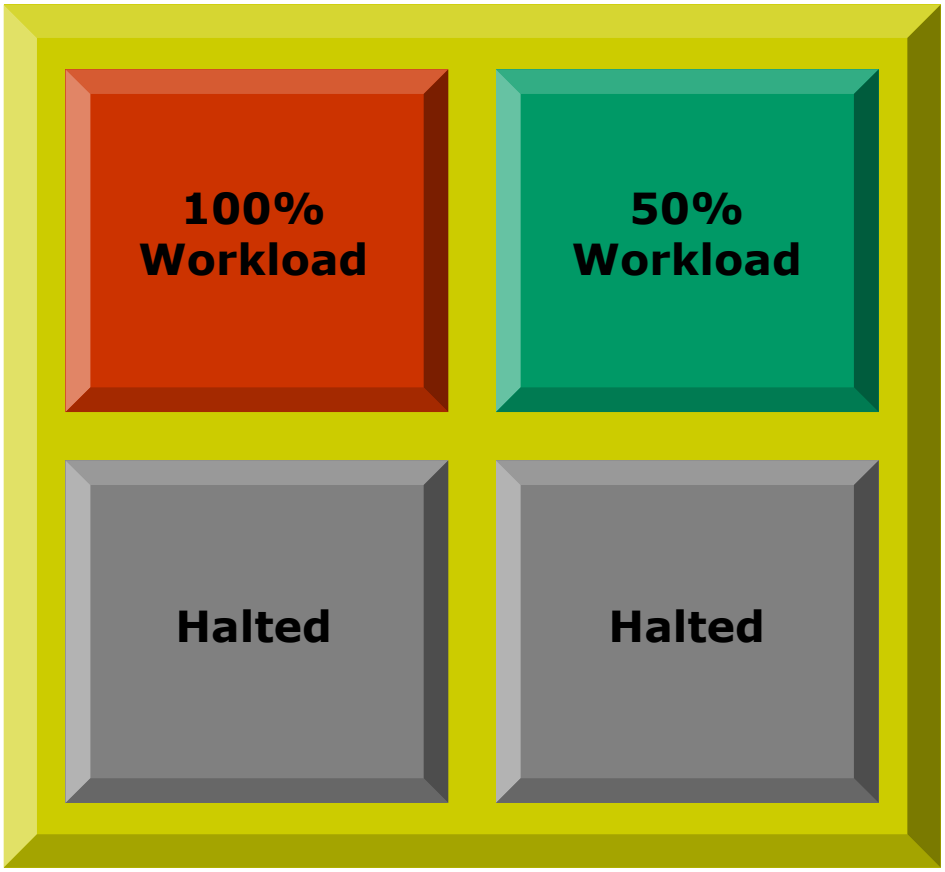
Ability to dynamically and individually adjust core frequencies to improve power efficiency

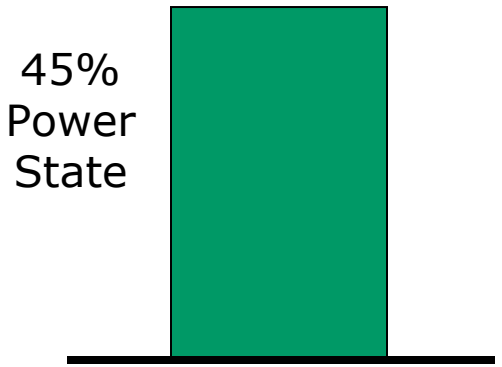




DICE: Dynamic Independent Core Engagement

Ability to dynamically and individually adjust core frequencies for improved power efficiency





Enjoy the rest of the conference !

www.amd.com/power

Trademark Attribution